

VERDIER Antonin

Diplôme de l'École Normale Supérieure (DENS)

08/07/2019 – 10/09/2019

--

2018/2019

École Normale Supérieure, Biology Department
PSL Université Paris

Statistical analysis of learning through behavioral experiments in mice

Perception et Mémoire / Décision et Processus Bayésiens–
Institut Pasteur

Supervisor: GUERINOT Corentin, Doctorant
Principal Investigators : LLEDO Pierre-Marie, MASSON Jean-Baptiste



Contents

1 Abstract	3
2 Introduction	4
2.1 Notation and Terminology	5
3 Materials and Methods	6
3.1 Data Acquisition	6
3.1.1 Calcium imaging	6
3.1.2 Behavioral Data	7
3.2 Data analysis	8
3.2.1 Preprocessing	8
3.2.2 Tensor Component Analysis	9
3.2.3 Importance of computational speed	11
3.2.4 TSNE	12
3.2.5 Random Forests Prediction	12
3.2.6 LSTM	12
4 Results	14
4.1 Factorplots	14
4.2 Random Forests	14
4.3 TSNE	16
4.4 Hierarchical Clustering	16
4.5 LSTM	17
5 Discussion	18
6 Conclusion	19
7 Acknowledgments	20
8 Appendices	21

1 Abstract

Learning is a fundamental phenomenon which occurs within neural networks and strongly impacts behavioral responses. During olfactory associative learning, the olfactory bulb plays a major role in early processing. In mice, recruitment of adult-born interneurons is a supplementary mechanism underlying plasticity. However, it is still unknown whether a statistical signature of learning can be extracted from calcium activity imaging of these neurons. Here, we explore with a new application of matrix decomposition methods the evolution of network activity through the learning process. Remarkably, our models succeeded at drastically reducing data dimension while keeping relevant learning information. Moreover, we were able to predict behavior based on this reduced data with high accuracy. Finally, we made the first steps into inter-animal comparison as we predicted behavior with information partially extracted from different animals. Taken together, our results suggest that adult-born interneurons calcium activity carry a statistical signature of learning, which could be similar across animals.

2 Introduction

Learning and memory are two major aspects of neuroscience nowadays. As imaging techniques improve and enable to investigate more and more complex structures, there is an increasing need for tailored algorithms suited for extracting understandable data. Indeed, the amount of data retrieved from experiments is unfortunately not often correlated with the amount of knowledge. Important efforts must be made to find appropriate spaces to visualize and interpret data.

The olfactory bulb (OB) is the first relay for odor information in a mouse's brain. Sensory neurons project their information to relay neurons, mitral and tufted cells (Figure 1), involved in early processing. Previous studies have shown reward-associated signals after associative learning in structures downstream of the OB [1][2]. Moreover, associative learning task modifies the firing activity of mitral and tufted cells.

Granule cells are inhibitory interneurons with a small cell body (6-8 μm). They have an axonless morphology and a dendritic output mainly relying on dendrodendritic synapses. Their connections are essentially made with mitral and tufted cells. They also receive a top-down input from the central brain. One important particularity of the OB is the presence of adult neurogenesis represented by adult-born granule cells that integrate themselves into the existing network. Therefore, we talk about structural plasticity, a mechanism acting in parallel of conventional functional plasticity. This has been shown to boost odor-reward association [3].

Adult-born GCs are also involved in various behavioral responses, such as odorant detection [4] and discrimination [5] and olfactory innate behaviors [6] [7] [8]. Consequently, adult-born GCs form an interesting population for the study of learning processes as they encapsulate numerous learning mechanisms, are accessible to imagery and play a role in both odor processing and behavioral response.

Several mice were submitted to reward-associated odor task during which calcium imaging was performed in their olfactory bulb through a cranial window. The neuronal activity extracted from this imaging, associated with the results of the behavioral experiments should allow us to link activity patterns with behavior.

Our main goal in this report is to explore new ways of extracting data from behavioral learning experiments. The main difficulty lies in the visualization of the evolution of such high-dimensional set of data. Hence, using dimensionality reduction methods, we aim to extract the statistical signature of learning from raw activity data associated with behavioral experiments. Each recording was made inside the olfactory bulb, at a constant depth, aiming to record dendrites of adult-born granule cells.

This report is part of a Ph.D. project which aims to explore novel statistical ways to extract learning information from brain activity data. The goal of this traineeship was to program several novelties analyzing methods, from standard Principal Component Analysis (PCA) to more complex neural networks such as Long Short-Term Memory (LSTM). All this work was done using Python programming language.

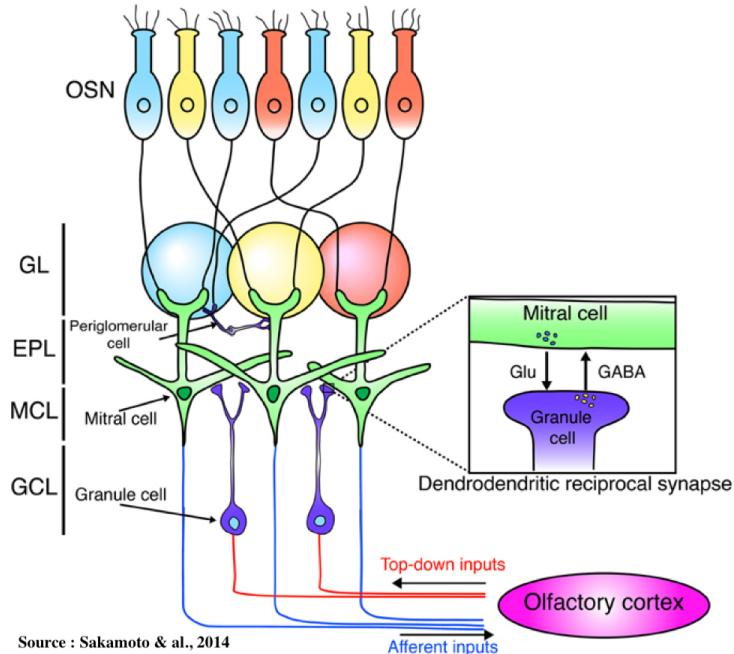


Figure 1: Basic model of the olfactory bulb network. Olfactory Sensory Neurons (OSN) makes synapses with the Glomeruli Layer (GL). Periglomerular cells (PG) surround the glomeruli. Both tufted cells (TC) and mitral cells (MC) in external plexiform layer (EPL) and mitral cell layer (MCL) send their output to the olfactory cortex. Granule cells in the granule cell layer (GCL) receive inputs from the olfactory cortex. They make synapses with mitral cells through dendro-dendritic synapses [9].

2.1 Notation and Terminology

Throughout this report, we will consider that a tensor is an array with more than 2 axes (or modes) and that a matrix is a two-dimensional array (2D array). Data will be represented as a 3D tensor D of shape $N \times T \times K$ where N is the number of region of interest (ROI)(i.e fragments of dendrites), T the number of timepoints at which activity was recorded and K the number of trials. Each element x of this tensor can be indexed by x_{ntk} with n, t, k ranging from 1 to N, T , and K respectively.

Tensors and matrices will be represented as uppercase letters whereas values from this tensors/matrices will be denoted by lowercase letters.

3 Materials and Methods

3.1 Data Acquisition

3.1.1 Calcium imaging

Lentivirus expressing the red fluorescent protein TdTomato and cre-recombinase combined with virus expressing floxed fluorescent calcium indicator (GCaMP6f) were injected into the subventricular zone (SVZ) to label adult progenitors (Figure 2A). These cells migrated to the olfactory bulb having red fluorescent structural labeling and calcium indicator in the green channel (Figure 2B). A glass cranial window was made over the olfactory bulb (Figure 2C-D). The mice was then trained to be head restrained and placed in a custom stage for stable, awake imaging under a 2-photon microscope (Figure 2E). To track the development of adult-born cells, the window surgery is performed first and a virus is injected later as outlined in Figure 2F. The new neuron structure (red) and neuronal activity (green) is imaged from their first arrival into the olfactory bulb until 2 months after virus injection. Both spontaneous and odor evoked responses are measured. The complete neuron structure is traced (red channel) and tracked with the development of the neurons.

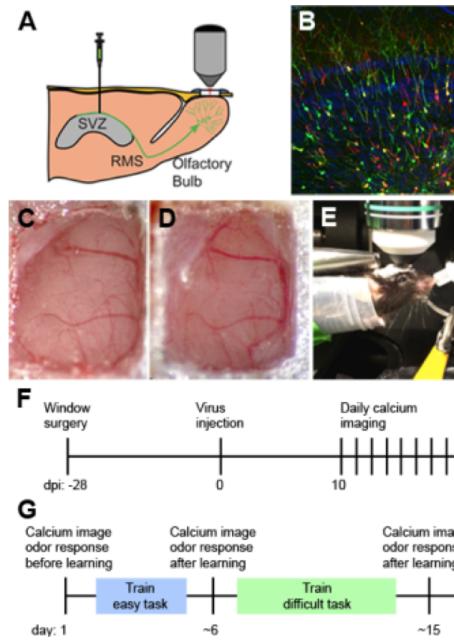


Figure 2: Labeling, cranial window and imaging of olfactory bulb adult neurogenesis. A) Virus injection into the subventricular zone (SVZ) where cells migrate via the rostral migratory stream (SVZ) to become interneurons in the olfactory bulb; 2-photon imaging through a cranial window. B) Labeling of adult-born neurons for structure (red: TdTomato) and function (green: GCaMP6f). C) Chronic cranial window at 1 month and D) 2 months. E) head-fixed imaging setup with odor delivery (white tube) and water reward (silver tube).

The first challenge for processing this raw data was to label Regions of Interest (ROI) where fluorescence signals related to calcium activity could be observed. Unfortunately, image label-

ing is a very complicated process to automate. Therefore, prior work has been done and all recordings were labeled by hand (Figure 3).

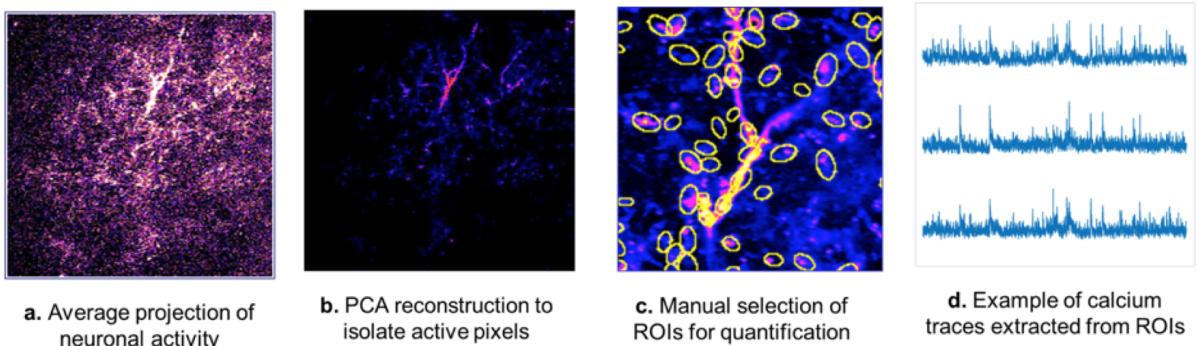


Figure 3: Early processing pipeline of calcium imaging data

Huge efforts were made to keep the imaging region constant for each mouse.

3.1.2 Behavioral Data

The learning process will be studied via an odor-reward association task subdivided into trials. During a trial, two odors can be presented, but one only is associated with a reward. If odor A is the reward-associated odor, the mouse, deprived in water, will get a reward when licking the water dispenser. If odor B is presented, no reward will be supplied. At the end of the learning process, the mouse should expect a reward only for the reward associated odor. For one trial, the recording is done for 20s at a 15Hz frequency, giving 300 time points.

Considering odor A as the reward-associated odor, four behaviors are possible. If the odor presented is A and if the mouse does lick the water dispenser, then the outcome is a HIT; if not, it is a MISS. If the odor presented is B and the mouse reach for the water dispenser, it is a FA (False Alarm). If not, it is a CR (Correct Rejection). Considering that the mice are deprived of water, CRs are the best indicators of learning because they do not bother to lick the water dispenser, knowing that nothing will come out.

During the whole learning experiments, many variants of associative learning can be performed. In this report, we mainly consider Easy Task 1, Easy Task 2 and Difficult Task 1, each one differs from another by the couple of odors presented. Easy Task 1 odors are Valdehyde and Cineole. Easy Task 2 odors are Hexanone and Ethyl Tigate and Difficult Task 1 two odors mixed with 40/60 % ratio and 60/40 % ratio. Trials are grouped in blocks of 20 trials. In each block, 10 reward-associated trials and 10 non reward-associated trials are randomly performed. Often, several blocks are conducted on the same day. For each block, a learning score is assessed as

follows :

$$\text{Learning score} = \frac{\#CR + \#HIT}{\#Trials} \quad (1)$$

This way, the performance of the mouse can be tracked over blocks and over days.

3.2 Data analysis

3.2.1 Preprocessing

The mouse is head-fixed under the microscope, however, this does not prevent motion of the field of view, especially during odor presentation or reward distribution where the mouse is prone to move. Frames (or time points) associated with aberrant fluorescent values due to motion are automatically replaced by NaNs (Not-a-Number values). If the trial does not meet preceding described criteria, data will be interpolated. In other words, blanks left by NaNs will be replaced by values linearly interpolated from the last and next known time points.

Raw activity data need to be preprocessed in order to remove all unrealistic and incomplete samples. In our dataset, around 15% of the data was classified as NaN (Not a Number). These non-values cannot be handled by later stages analysis so they need to be removed or replaced. If a trial has more than 80 NaNs in total, it is removed from the dataset. If a trial has a series of at least 25 consecutive NaNs, it is also removed from the dataset. These are thresholds chose to preserve data integrity.

Raw data also contains some artifacts, mainly related to the cranial window narrowing due to bone reconstruction. In these cases, baseline fluorescence F_0 was too low and therefore normalized fluorescence fluctuations were extremely high. Corresponding trials/ROIs were removed empirically when the normalized fluorescence $\frac{df}{f}$ range exceeded 15.

After NaNs removal, we produce three datasets of the activity. The first one is raw, the second one is normalized and the third is normalized and smoothed. Normalization can be done on all the data, across trial or depending on the day the experiment was performed (normalization will be done across trials conducted the same day). Normalized data will then be smoothed following a Univariate Spline method. All three preprocessed datasets will be saved for later use. All these steps are integrated into an automated pipeline (Figure 4). It will take several minutes to achieve complete pre-processing for the data of one animal, considering a tensor of shape (260, 285, 315).

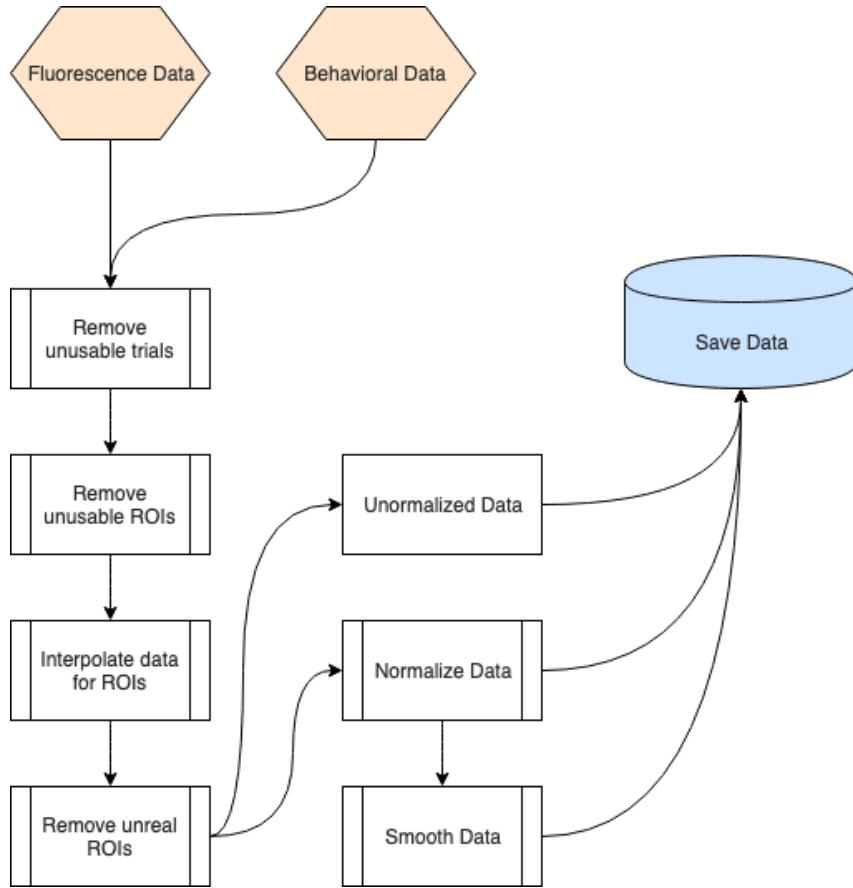


Figure 4: Flowchart of the preprocessing pipeline. Data is given to the pipeline on an animal basis

3.2.2 Tensor Component Analysis

In order to describe TCA (Tensor Component Analysis) and its usefulness, we first need to describe PCA [10] (Principal Component Analysis).

Considering a tensor D (as defined in Notation and Terminology) of N neurons during T time points over K trial experiments, data will be represented by a tensor of shape $N \times T \times K$. Each element in this tensor coincide to the fluorescence ratio for a given ROI, at a given time point for a given trial. The main problem is that such large dataset is difficult to interpret as is. One common workaround is to consider inter-trial differences as noise that need to be reduced and therefore averaging data across trials. This leads to a matrix of shape $N \times T$ with elements x_{nt} . PCA is then performed to extract r components explaining most of the dataset variance :

$$\bar{x}_{nt} \approx \sum_{r=1}^R w_n^r b_t^r \quad (2)$$

This outputs two factors matrices W and B , of shape $N \times r$ and $T \times r$ respectively. The neuron factor matrix, W , will describe with r component the behavior across neurons and the time factor matrix, B , across time points. These factors can only be described this way if the

input matrix has a shape of $N \times T$ and this why data needs to be averaged across trials.

TCA (Tensor Component Analysis) is a novel approach for analyzing neuronal data proposed by Williams & al. [11]. It is also based on matrix decomposition, a well-known mathematical field where the matrix data is expressed as a product of matrices. It is a non-supervised method. TCA is similar to PCA in a way that it extracts abstract features supposed to be representative of the overall data (Figure 5). With the same formalism as before, we can define TCA as :

$$\bar{x}_{ntk} \approx \sum_{r=1}^R w_n^r b_t^r a_k^r \quad (3)$$

Where k denotes a trial. This way, TCA outputs three factors W , B and A that can be described as neuronal, time and trial factors respectively. The advantages of this method are, of course, to benefit from this third trial factor of shape $K \times r$ that encapsulates dynamics across trials.

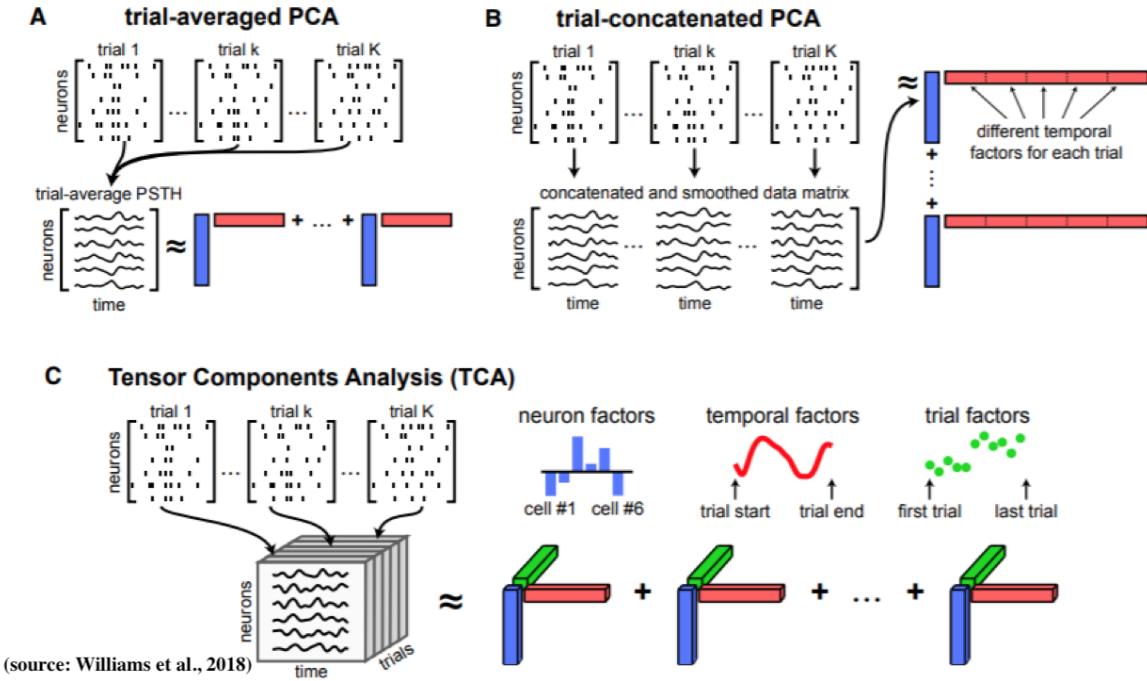


Figure 5: Figure A. Trial averaged PCA method. Trial specificity is crushed into one dimension to allow computation. This method does not encapsulate evolution across trials. Figure B. Trial concatenated PCA. Trials data are concatenated one after another. This allows integrating trial evolution into time factor. Figure C. Tensor Component Analysis. This method keeps trial evolution information and is able to process it independently from other factors.

TCA also gets another advantage over PCA from the fact that these factors can be non-orthogonal, whereas PCA only allows orthogonal factors to be solution. In other words, it states that PCA components are not correlated one to another. Consequently, this means that TCA is much more likely to fit real-world neuronal data, where components can be correlated. By defi-

nition, PCA solution is unique but TCA can have several local solutions, therefore optimization is more difficult and constitutes a non-convex problem. Thereby we must start with initialization and proceed with optimization algorithms. Finally, PCA components can be ordered by variance explained, in contrary to TCA.

Initialization of the TCA factors could be done according to two techniques. The first one, more efficient for convergence is Singular Value Decomposition (SVD). However it may influence the final results and the chance to reach a local minimum is greater than with the second method, a standard random initialization. Random initialization was mainly used for these simulations, even if it is computationally costlier.

In this report, we investigate three types of TCA to see if any of them better fit the data and allow more precise interpretation. We will define these types as Non-Negative Parafac, Regular Parafac and Custom Parafac. The 'Parafac' term refers to the model used for the tensor decomposition. Non-Negative Parafac (NNP) doesn't support negative factors (i.e $w, b, a > 0, \forall w, b, a$). Regular Parafac (RP) allows factors to be negative or positive. Finally, Custom Parafac (CP) only allows one factor to be negative.

We also experimented a completely new method that we called Fixed Parafac. The goal behind it is to perform data analysis not only with a unique model for each animal but with a model for several animals. To do this, time factors were extracted from NNP performed on 3 different animals. We then rearranged these sets of 6 time factors to obtain 6 groups of time factors as homogenous as possible. Finally, each group was averaged to get one general time factor for each of the 6 components. When performing Fixed Parafac, time factor is not allowed to be optimized and is replaced with this general set of inter-animal averaged time factors. This way, TCA will be allowed only to optimize neuron factor and trial factor and therefore will be forced to fit animal-related data taking into account inter-animal data.

ALS (Alternating Least Squares) is the method used for the optimization of these decompositions. The three-factor matrices are optimized each after another by fixing two of them and optimizing the third.

3.2.3 Importance of computational speed

One main aspect when evaluating ways to analyze data is computation speed. If a model has 10 hyperparameters and takes 2 hours to run each time, its optimization will be a long process and eyeballing hyperparameters will often be a solution. A lot of work has been put in prior speed optimization of the model in order to be able to optimize sharply its hyperparameters in

a second time. Originally, TCA optimization would take around 0.5s for each iteration, for an average 1000 to 2000 iteration before reaching convergence threshold. Thanks to the use of GPU parallelization we brought the time for an iteration to 0.04s. After some more complex, python-related optimization, an average of 0.029s per iteration was reached with a high-end customer Graphic Processing Unit (GPU).

3.2.4 TSNE

T-SNE (T-Distributed Stochastic Neighbor Embedding) is a dimension reduction technique for data visualization. Given high-dimensional data, it outputs 2D or 3D data in a scatter plot fashion. It is a non-linear technique preserving the notion of distance between points while warping space.

3.2.5 Random Forests Prediction

Random forests classifiers are part of the supervised learning models. A random forest is composed of numerous decision trees (Figure 6, each one is trained on a subset of the data). A decision tree, given a set of data, will try to split this data according to criteria. These criteria must split the data as efficiently as possible, meaning that with a simple condition, the two output groups are far apart. To measure the quality of a split, Gini impurity was used.

When all the trees are trained on their data, a prediction for a new input can be made. Each tree will output a value based on its training experience. The value which gathers most votes among others will be the final output of the forest. This method is known for being very robust against over-fitting (when the model is too close to the data and thus cannot be correctly generalized). One major asset for random forests is that we can extract feature importances, meaning that we have an idea of which feature is the most predictive. The predictive power of each feature is given in percentages.

3.2.6 LSTM

LSTM networks are neural networks. Neural networks are neurons in layers, each layer being connected with the previous one. The neuron gets inputs from all neurons in the previous layer, each input being weighted differently. Updating these weights is a way for the network to learn. LSTM are special networks in a sense that they can remember previous inputs. This particularity is interesting here as learning is a process happening through time. To have this

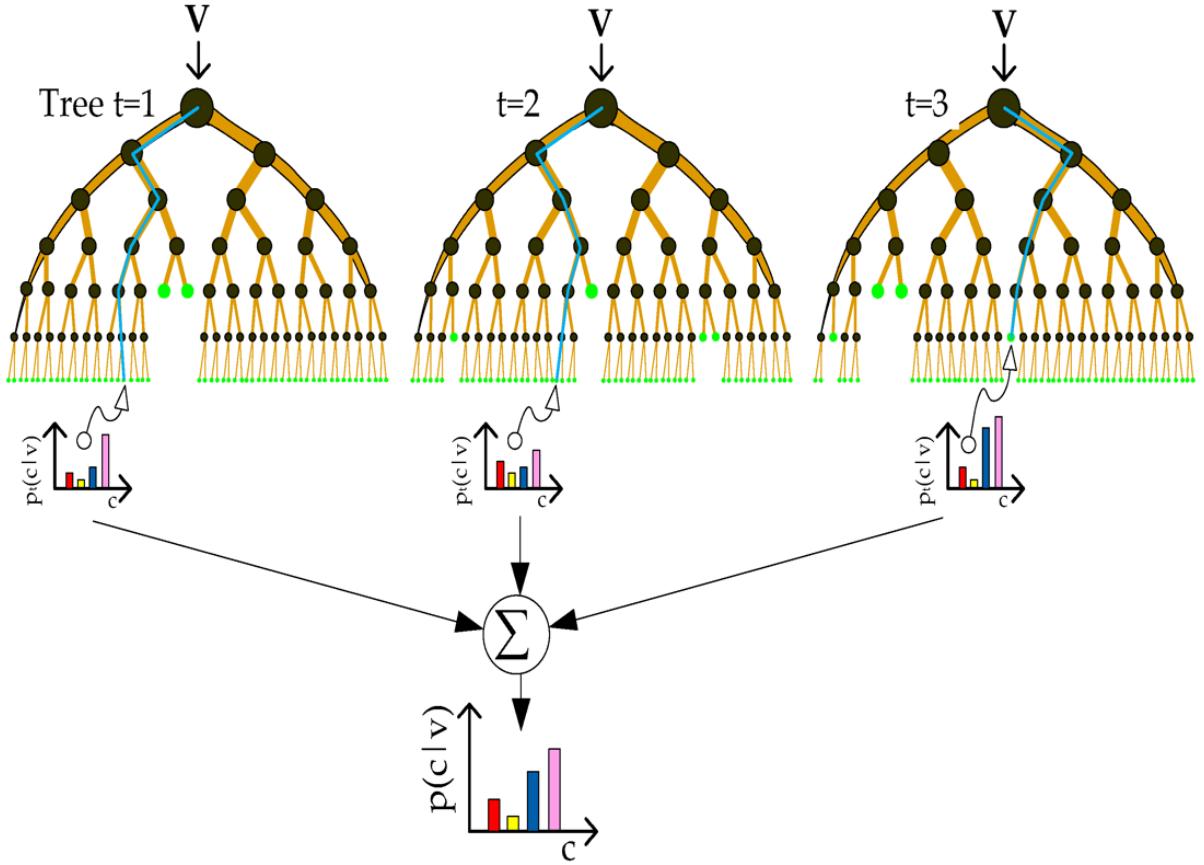


Figure 6: Random forest scheme. Each decision tree is called an "estimator". Based on intrinsic criteria, a tree will vote for a specific output. The output which gets the maximum score over all estimators is the final output of the forest.

capacity, the basic neuron of an LSTM network has three gates : a forget gate, an input gate and a sigmoid gate (Figure 7).

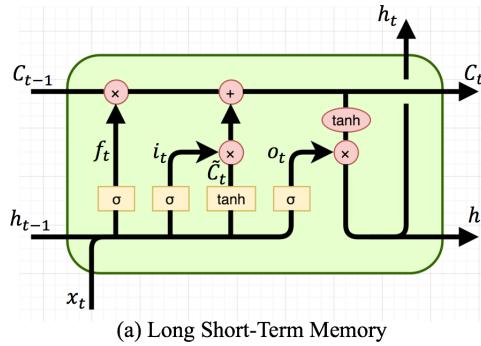


Figure 7: LSTM neuron scheme

It is important to differentiate the "Cell State" (CS) from the "Hidden State" (HS), two basic features in an LSTM network. The CS, C_t on the figure, is specific to this kind of network. It carries a flow of information across all the network. The HS, H_s on the figure, represents the real output of the neuron, as in other networks.

The forget gate, f_t , is responsible for keeping or eliminating information across the network.

The input gate, i_t , choose to add new information to the CS. Finally, the sigmoid gate, o_t , is the output information. Together, these three gates make the gist of this kind of network.

4 Results

4.1 Factorplots

Considering the TCA method described before, we can analyze our data to extract three factors : a neuronal factor, a time factor and a trial factor (Figure 8). 6 components were extracted from the TCA because it appears to be a good balance between information stored and computational cost (Figure ??). Neuron factors are represented according to their spatial configuration thanks to a manually generated mask (i.e each ROI was surrounded by hand) unique to each animal. However, this represents a slice that can be perpendicular to dendrites of GC therefore two close ROIs do not necessarily belong to the same GC.

In this figure, we can see that the neuron factor could reflect a mass activity (component 4) or a more localized one (component 1 or 5). Time factor can be subdivided into odor-evoked (component 1 or 3) and reward-evoked (component 2). Finally, for some components, group are distinguishable for components of the trial factor. Consequently, it seems likely that information related to the odor and/or the reward are extracted inside these components, and that it could be used for predicting behavior of the mouse. Results were similar across the 5 parafac types.

4.2 Random Forests

After performing TCA on neuronal activity, we want to check whether or not the components from TCA contain relevant information regarding the experiment. Using Random Forest, we predicted from trial factors which odor was smelt by the mouse, and if the mouse did take the reward. We succeeded around 95% accuracy for odor prediction and 90% accuracy for reward prediction on several animals for easy task and difficult task protocols. As said before, we mainly focused on 6-component TCA, as odor prediction and reward prediction scores did not increase significantly with more components (Figure 9).

As explained earlier, random forests provide a huge advantage to rank each feature by predictive importance. We were therefore able to rank each component of the TCA according to their predictive importance. This gave us a first insight of which ROIs were most involved in odor prediction. To see if the information related to an odor is focalized on several neurons, we looked at most the predictive ROIs for a given odor (Figure 10).

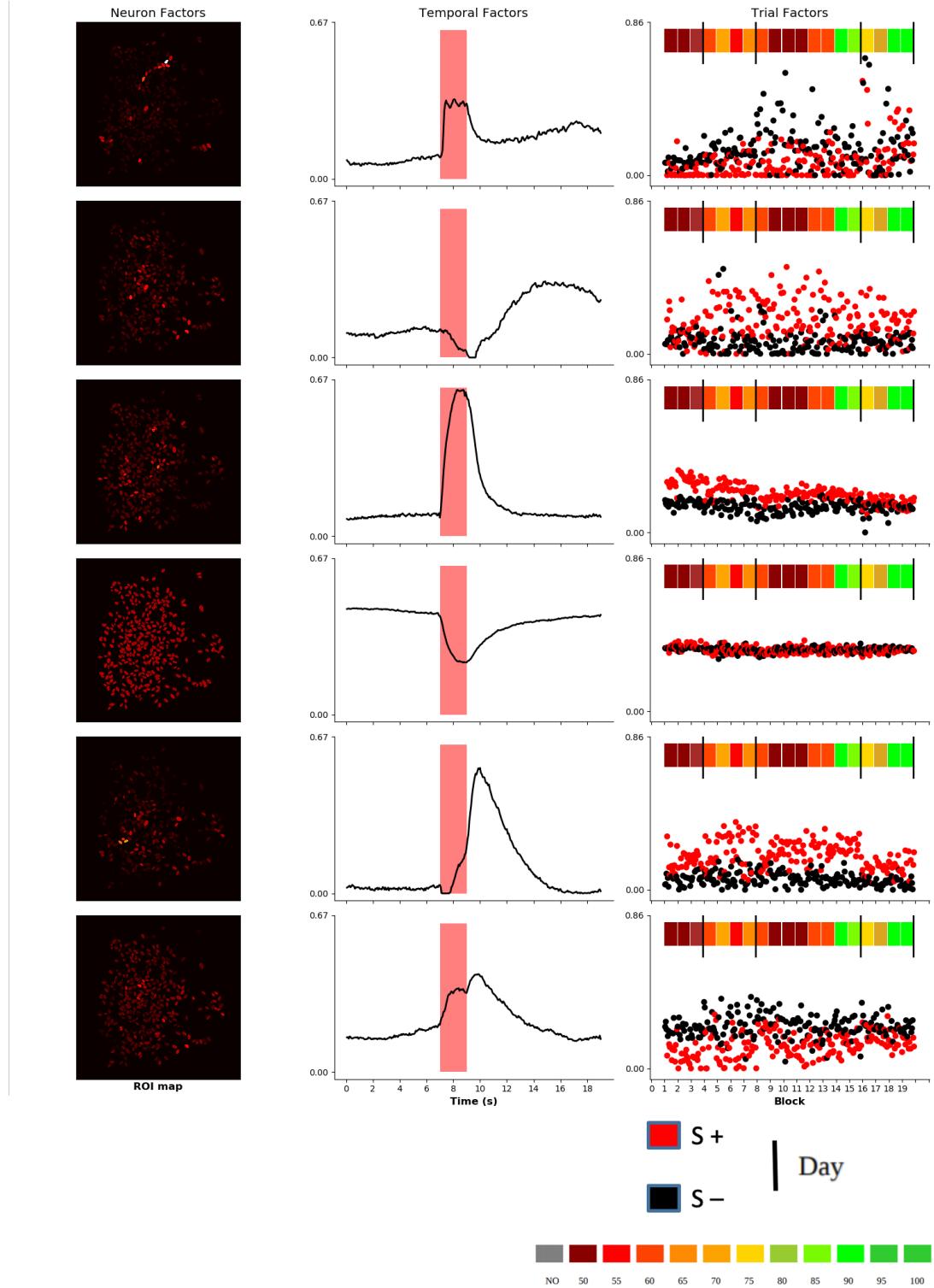


Figure 8: Representation of the factors of a non-negative-parafac. Neuron factors (on the left) are shown in their biological spatial configuration. Presentation of the odor (for 2s) is shown by the red rectangle on the time factor plot (on the middle). Finally, trial factors (on the right) are colored according to the odor presented (red corresponds to the reward-associated odor and black to the non reward-associated odor). Block overall performance is illustrated by the color of the rectangles on the top of the trial factor plot. Red blocks correspond to approximately 50% of correct answer opposed to green block (95%).

As we can see, for different odors, the information seems to be processed in different ROIs (i.e different neurons).

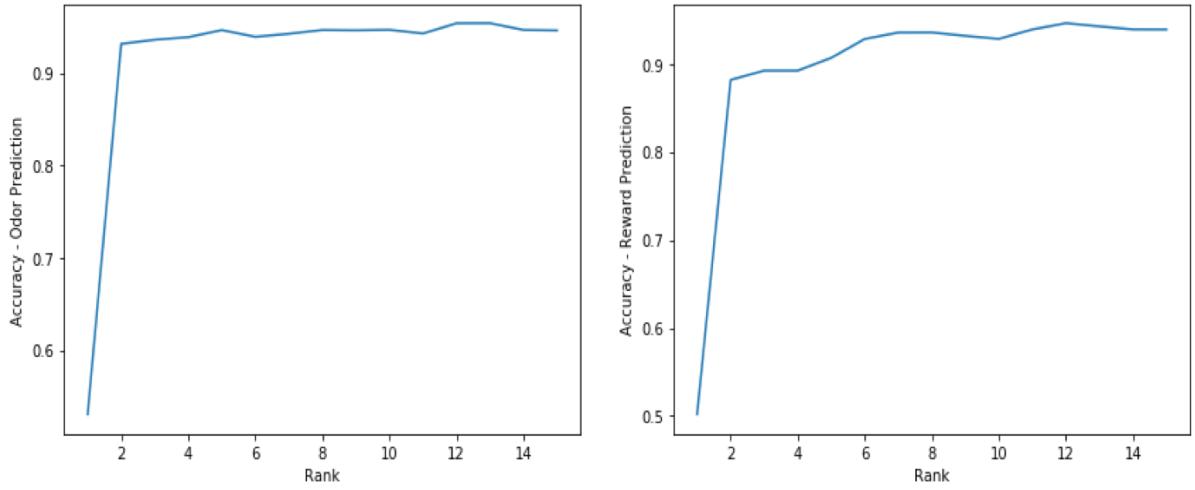


Figure 9: Prediction accuracy depending of the number of component extractes from the TCA.

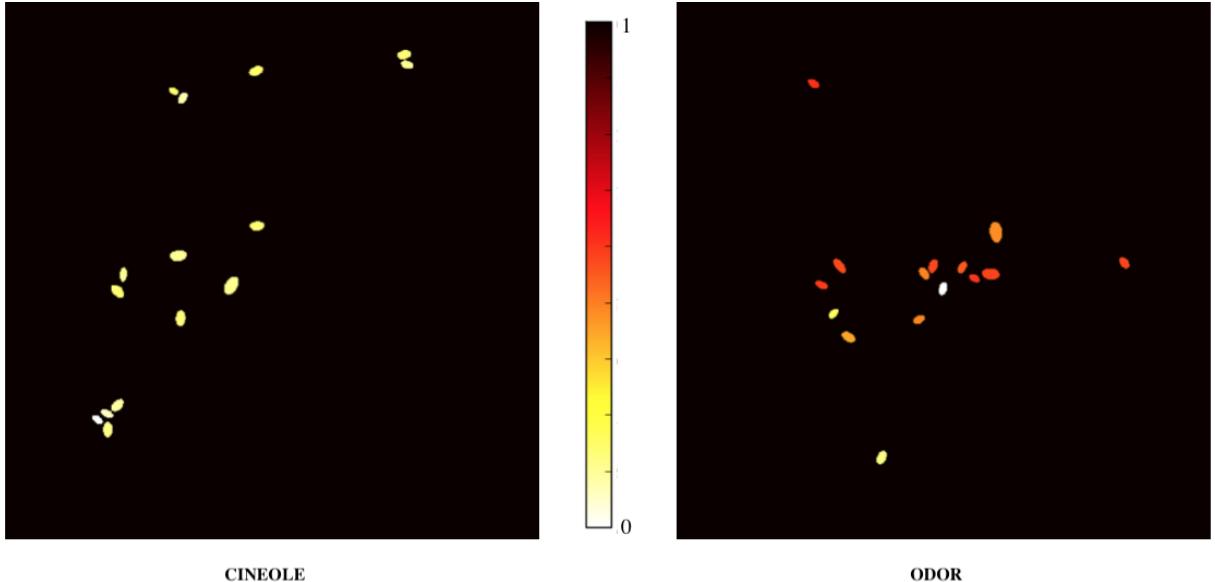


Figure 10: Maps of most predictive ROIs for VALDEHYDE and CINEOLE, colored by normalized activity, for the same animal.

4.3 TSNE

After the subtle approach of TCA which aims at extracting interpretable factors from raw data, more computationally expensive methods were tried as non-supervised clustering algorithms like TSNE (Figure 11).

4.4 Hierarchical Clustering

According to the fact that our model predicted with 90% accuracy behavior, we attempted to see the evolution of factors through trials to find out if we can distinguish anything related to learning. In other words, does the neuron factor aspect evolves with learning performance (Figure 12). Hierarchical clustering of the correlation matrix was performed for each block for

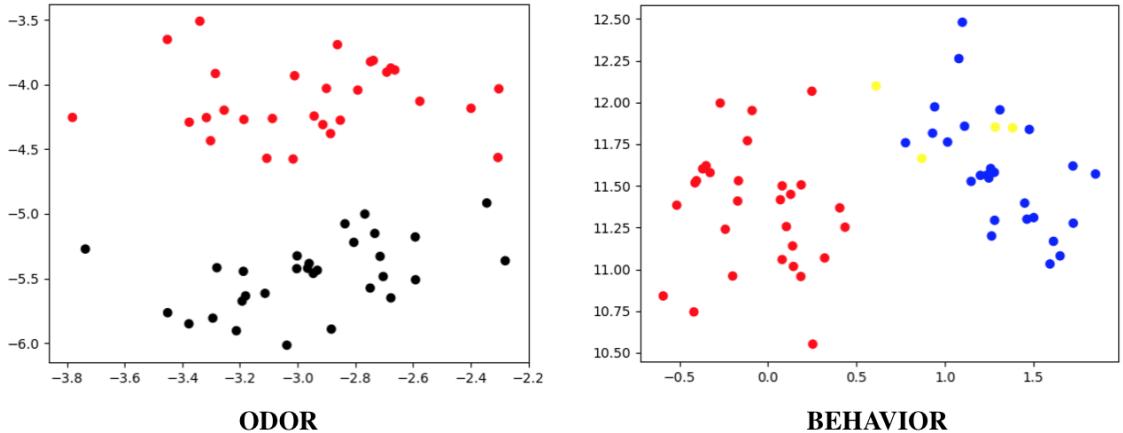


Figure 11: TSNE representation of Easy Task 1 trials. Left panel is colored according to the odor presented. Red is the reward associated odor and black is the non-reward associated odor. Right panel is colored according to the response behavior of the mouse. HIT trials are colored in red, CR in blue and FA in yellow.

a day. Only trials with similar behavioral response were analyzed together. We did not succeed at seeing a link between the performance of the mice and the aspect of the heatmap.

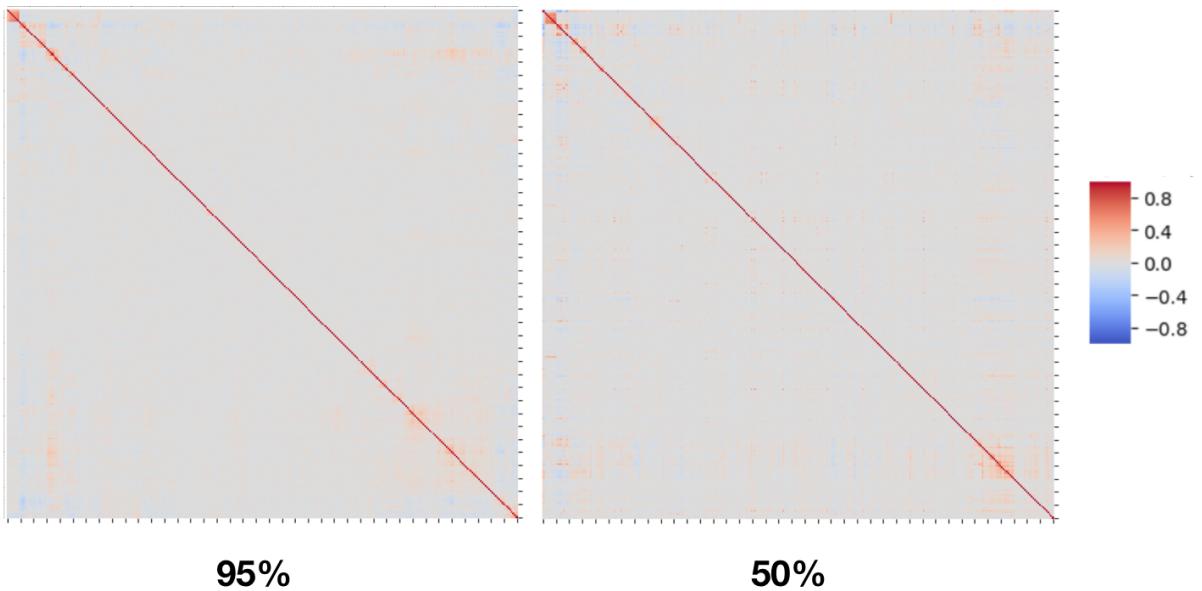


Figure 12: Hierarchical clustering representation of ROIs. The left panel is a hierarchical clustering of the mean correlation matrix of CR trials from one block with high performance (95% learning score). The right panel is a hierarchical clustering of the mean correlation matrix of CR trials from one block with high low performance (50% learning score). These two blocks were performed the same day on the very same animal. Right panel ROI order were aligned on left panel hierarchical clustering order for comparison. No clear difference can be seen in these figures.

4.5 LSTM

LSTM networks are the gold-standard prediction tool for time-related data. A neural network (Figure 13) was trained on raw calcium-related activity to predict the next behavior of the mouse

given all precedent ones. 100% accuracy was reported for odor classification and 95% accuracy for reward classification.

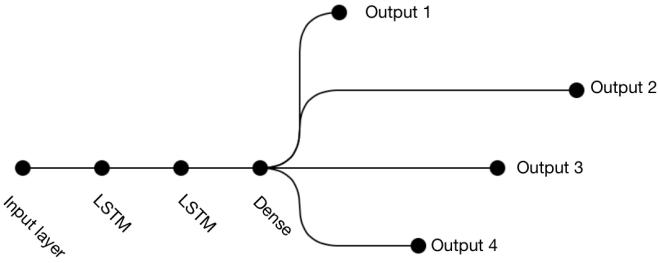


Figure 13: Neural network structure used to predict next behavior. The first layer takes an activity tensor as an input. Two consecutive layers of 32 and 16 LSTM cells respectively are used as the core predictors unit of the network. A dense layer of 16 formal cells then projects to 4 distinct output neurons with a sigmoid activation function.

5 Discussion

Tensor Component Analysis has demonstrated its capacity to extract interpretable factors from abstract raw calcium-related activity data. Temporal factors seem very helpful for understanding the role of a given component. As an example, components with high predictive importance for odor tend to have a time factor with a quick rise reacting at the beginning of the odor presentation. Besides, reward-associated components rise quickly at the end of the presentation, when the reward must be taken or not. After dozens of simulations, we were capable of making groups of time factors curves, without much leftovers. This can mean that time-factor possibilities are limited. Moreover, fixed parafac and fixed non negative parafac had similar results than other types of TCA, making inter-animal comparison possible. This approach was completely new and considering its first promising results, it should be further investigated. Interestingly, a simple linear decomposition method such as TCA enables to retrieve trial-associated information. This could be reflecting an actual linear integration of inputs in the olfactory bulb[12].

The difference between regular parafac and non-negative parafac did not appear as relevant. By allowing the neuron factor to be negative, the model would shift the first component from a positive value to a negative value and therefore affect the overall intensity but not dynamics. Custom parafac behaved the same as regular parafac as one of the component would absorb all negative values.

Furthermore, trial factors, the main addition permitted by TCA, appear to be a good indicator on its own regarding discrimination between odors. On most predictive components for

reward or odors, trial factors were often far apart, meaning that the distance between the trials was correctly extracted from raw activity. Associated with neuron and time factors, they help understand which components were more important for odor A or odor B or for reward.

We are imaging a very small subset of GCs, which are not the principal neurons for information processing in the olfactory bulb [13]. Thus, it suggests that information is very redundant in the olfactory bulb. This kind of redundancy ensures robustness of the network and prevents loss of information.

One major drawback of the TCA is the sensitivity to the spatialization of the ROIs. As ROIs are manually selected on an animal basis, it exists an intrinsic limitation preventing inter-animal comparison. We are partially addressing this problem with the fixed parafac method, where the time factor reflects TCA output for several animals.

Finally, we have to keep in mind that we are imaging calcium-related fluorescence and not voltage. Hence, we could miss information on more subtle dynamics.

As we located in the olfactory bulb, it is not unreal to see 95% accuracy for odor prediction. Indeed, it has been shown that early treatment of the information was made, however, the biological support for this information (i.e odor pattern representation) remains unclear. It is much more interesting to see information about reward regarding how early the olfactory bulb is in the information chain. Surely LSTM performed a little better due to their capacity to retain information across inputs but do not give any insights on neural activity dynamics as TCA does.

6 Conclusion

We investigated a new approach for analyzing neuronal activity data by predicting mouse behavior from raw activity with Tensor Component Analysis. Predictions results obtained with this method were satisfying. More optimization of TCA's hyperparameters could give us more accuracy. This type of analysis has the great benefit to produce understandable factors, that can be interpreted easily. Moreover, it allows one to have an insight on the evolution of activity across trials, which is usually not the case with traditional methods. Fixed parafac and fixed non-negative parafac are a first step into inter-animal comparison which is a major limitation in this field.

The statistical signature of learning is present in final factors as it is possible to predict the behavior of the mouse based on previous trials. However, more work should be made to extract this particular information. A "score" of instantaneous learning could be associated with a

particular brain activity however this score cannot be absolute and must be integrated into the learning process of a particular animal.

As a conclusion to this report, it is important to emphasize the encouraging results obtained with these methods. The importance of discovering a way to measure the learning performance of a network from real-world data is such that investigations must continue. Supplementary pseudo-learning experiments are scheduled for comparison purposes.

7 Acknowledgments

I especially thank Corentin Guerinot for his great supervising, presence and critical reading of this report. I also thank Kurt Sailor for his help with imaging protocol description and critical reading of this report. I finally thank the labs Perception and Memory, Decision and Bayesian Computation for their warm welcome.

References

- [1] Matthew R. Roesch, Thomas A. Stalnaker, and Geoffrey Schoenbaum. Associative encoding in anterior piriform cortex versus orbitofrontal cortex during odor discrimination and reversal learning. *Cerebral Cortex (New York, N.Y.: 1991)*, 17(3):643–652, March 2007.
- [2] David H. Gire, Jennifer D. Whitesell, Wilder Doucette, and Diego Restrepo. Information for decision-making and stimulus identification is multiplexed in sensory cortex. *Nature Neuroscience*, 16(8):991–993, August 2013.
- [3] Anne Grelat, Laura Benoit, Sébastien Wagner, Carine Moignau, Pierre-Marie Lledo, and Mariana Alonso. Adult-born neurons boost odor-reward association. *Proceedings of the National Academy of Sciences of the United States of America*, 115(10):2514–2519, 2018.
- [4] Vincent Breton-Provencher, Morgane Lemasson, Modesto R. Peralta, and Armen Saghatelyan. Interneurons produced in adulthood are required for the normal functioning of the olfactory bulb network and for the execution of selected olfactory behaviors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(48):15245–15257, December 2009.
- [5] Emeka Enwere, Tetsuro Shingo, Christopher Gregg, Hirokazu Fujikawa, Shigeki Ohta, and Samuel Weiss. Aging results in reduced epidermal growth factor receptor signaling, diminished olfactory neurogenesis, and deficits in fine olfactory discrimination. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 24(38):8354–8365, September 2004.
- [6] Masayuki Sakamoto, Itaru Imayoshi, Toshiyuki Ohtsuka, Masahiro Yamaguchi, Kensaku Mori, and Ryoichiro Kageyama. Continuous neurogenesis in the adult forebrain is required for innate olfactory responses. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20):8479–8484, May 2011.

- [7] Lillian Garrett, Jingzhong Zhang, Annemarie Zimprich, Kristina M. Niedermeier, Helmut Fuchs, Valerie Gailus-Durner, Martin Hrabě de Angelis, Daniela Vogt Weisenhorn, Wolfgang Wurst, and Sabine M. Höltner. Conditional Reduction of Adult Born Doublecortin-Positive Neurons Reversibly Impairs Selective Behaviors. *Frontiers in Behavioral Neuroscience*, 9, November 2015.
- [8] Nagendran Muthusamy, Xuying Zhang, Caroline A. Johnson, Prem N. Yadav, and H. Troy Ghashghaei. Developmentally defined forebrain circuits regulate appetitive and aversive olfactory learning. *Nature Neuroscience*, 20(1):20–23, January 2017.
- [9] Shin Nagayama, Ryota Homma, and Fumiaki Imamura. Neuronal organization of olfactory bulb circuits. *Frontiers in Neural Circuits*, 8, September 2014.
- [10] Karl Pearson F.R.S. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572, November 1901.
- [11] Alex H. Williams, Tony Hyun Kim, Forea Wang, Saurabh Vyas, Stephen I. Ryu, Krishna V. Shenoy, Mark Schnitzer, Tamara G. Kolda, and Surya Ganguli. Unsupervised Discovery of Demixed, Low-Dimensional Neural Dynamics across Multiple Timescales through Tensor Component Analysis. *Neuron*, 98(6):1099–1115.e8, 2018.
- [12] Priyanka Gupta, Dinu F. Albeanu, and Upinder S. Bhalla. Olfactory bulb coding of odors, mixtures and sniffs is a linear sum of odor time profiles. *Nature Neuroscience*, 18(2):272–281, February 2015.
- [13] Pierre-Marie Lledo and Matt Valley. Adult Olfactory Bulb Neurogenesis. *Cold Spring Harbor Perspectives in Biology*, 8(8), 2016.

8 Appendices

All code written during this internship is available here (2400 lines of code) : https://github.com/cguerino/pipeline_tca